

Claims

The invention claimed is:

Claim 1. In a distributed, multinode data processing system in which nodes communicate via adapters coupled to each node, a method for detecting node reachability inconsistencies in the presence of temporary communication failure or temporary daemon blockage, said method comprising:

upon detecting at a first node conditions suitable for instituting change in adapter group membership, sending to a recipient node, a message indicative of membership in a previous stable adapter membership group upon the condition that said recipient node is also indicated as being part of said stable group; and

initiating a group join protocol in which said message is rejected by a recipient node under a condition selected from the group of conditions consisting of (1) said message includes indicia showing that the sending node and the recipient node are in the same adapter membership class, said indicia being in conflict with corresponding indicia stored within said recipient node; and (2) said message includes indicia showing that the sending node and the recipient node are in different adapter membership classes, said indicia being in conflict with corresponding indicia stored within said recipient node.

Claim 2. The method of claim 1 in which said conditions include failure to detect heartbeat messages at said first node.

Claim 3. The method of claim 2 further including a step in which said first node, after its first message is rejected, transmits a second message which is indicative of nonmembership in a previously stable group, whereby the goal of having the sender and receiver agree on adapter group membership is advanced.

Claim 4. The method of claim 3 further including the step of processing said second message so as to provide mutually consistent views of the liveness of nodes in the membership group.

Claim 5. A multinode data processing system comprising:

a plurality of data processing nodes connected in a network capable of transmitting
5 messages between nodes;

storage means within said nodes containing program code for transmitting to a recipient node, upon detecting, at said first node, conditions suitable for instituting change in adapter group membership, a message indicative of membership in a previous stable adapter membership group upon the condition that said recipient node is also indicated as being part of said stable group and for initiating a group join protocol in which said message is rejected by a recipient node under a condition selected from the group of conditions consisting of (1) said message includes indicia showing that the sending node and the recipient node are in the same adapter membership class, said indicia being in conflict with corresponding indicia stored within said recipient node; and (2) said message node includes indicia showing that the sending node and the recipient node are in different adapter membership classes, said indicia being in conflict with corresponding indicia stored within said recipient node.

Claim 6. The data processing system of claim 5 in which said program code further includes means for transmitting from said first node, after its first message is rejected, a second message which is indicative of nonmembership in a previously stable group, whereby the goal of having
20 the sender and receiver agree on adapter group membership is advanced.

Claim 7. The data processing system of claim 6 in which said program code further includes means for processing said second message so as to provide mutually consistent views of the liveness of nodes in the membership group.

Claim 8. A program product comprising machine readable medium containing program code for use in a multinode data processing system for transmitting to a recipient node, upon detecting at said first node, conditions suitable for instituting change in adapter group membership, a message indicative of membership in a previous stable adapter membership group upon the condition that said recipient node is also indicated as being part of said stable group and for initiating a group join protocol in which said message is rejected by a recipient node under a condition selected from the group of conditions consisting of (1) said message includes indicia showing that the sending node and the recipient node are in the same adapter membership class, said indicia being in conflict with corresponding indicia stored within said recipient node; and (2) said message includes indicia showing that the sending node and the recipient node are in different adapter membership classes, said indicia being in conflict with corresponding indicia stored within said recipient node .

Claim 9. The program product of claim 8 further including means for transmitting from said first node, after its first message is rejected, a second message which is indicative of nonmembership in a previously stable group, whereby the goal of having the sender and receiver agree on adapter group membership is advanced.

Claim 10. The program product of claim 9 further including means for processing said second message so as to provide mutually consistent views of the liveness of nodes in the membership group.